





Clinical assessment of W-band spectroscopy for non-invasive detection and monitoring of sustained hyperglycemia

ALDO MORENO-OYERVIDES,^{1,2,*}  **M. CARMEN AGUILERA-MORILLO,³** **MARÍA JOSÉ DE LA CRUZ FERNÁNDEZ,²** **EDURNE LECUMBERRI PASCUAL,⁴** **LUCÍA LLANOS JIMÉNEZ,²** **VIKTOR KROZER,⁵** **AND PABLO ACEDO^{1,2}** 

¹*Department of Electronic Technology, Universidad Carlos III de Madrid, Leganés, 28911 Madrid, Spain*

²*Instituto de Investigación Sanitaria Fundación Jiménez Díaz (IIS-FJD), Madrid, Spain*

³*Department of Applied Statistics and Operational Research, and Quality, Universitat Politècnica de València, 46022 Valencia, Spain*

⁴*Servicio de Endocrinología y Nutrición, Hospital Universitario Ramón y Cajal, Madrid 28034, Spain*

⁵*Physics Institute, Goethe University Frankfurt am Main, Frankfurt am Main 60438, Germany*

*aldmoren@ing.uc3m.es

Abstract: HbA1c is the gold standard test for monitoring medium/long term glycemia conditions in diabetes care, which is a critical factor in reducing the risk of chronic diabetes complications. Current technologies for measuring HbA1c concentration are invasive and adequate assays are still limited to laboratory-based methods that are not widely available worldwide. The development of a non-invasive diagnostic tool for HbA1c concentration can lead to the decrease of the rate of undiagnosed cases and facilitate early detection in diabetes care. We present a preliminary validation diagnostic study of W-band spectroscopy for detection and monitoring of sustained hyperglycemia, using the HbA1c concentration as reference. A group of 20 patients with type 1 diabetes mellitus and 10 healthy subjects were non-invasively assessed at three different visits over a period of 7 months by a millimeter-wave spectrometer (transmission mode) operating across the full W-band. The relationship between the W-band spectral profile and the HbA1c concentration is studied using longitudinal and non-longitudinal functional data analysis methods. A potential blind discrimination between patients with or without diabetes is obtained, and more importantly, an excellent relation ($R\text{-squared} = 0.97$) between the non-invasive assessment and the HbA1c measure is achieved. Such results support that W-band spectroscopy has great potential for developing a non-invasive diagnostic tool for in-vivo HbA1c concentration monitoring in humans.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

The management of diabetes mellitus (DM) is a global concern due to its severe impact on health, its ever-increasing prevalence rate around the world, and other alarming socio-economic consequences [1]. This metabolic disorder is characterized by abnormal levels of glucose in blood caused by the inability of the human body to produce or use insulin adequately [2]. In this regard, it has been widely recognized that medium/long-term exposure to high-glucose concentration in blood (sustained hyperglycemia) leads to the irreversible formation of Advanced Glycation End-products (AGEs) [3–7], which play a key role in the development of several chronic DM complications. Consequently, a strict control of glycemia is critical in the management of DM to reduce the risk of developing serious health complications.

Nowadays, in addition to the standard measurement of the instantaneous Blood Glucose Level (BGL), medical proceedings for DM care use the HbA1c concentration (glycated hemoglobin)

as a valuable and reliable indicator of mean glycemia value over the past 2-3 months [8,9]. In fact, it is advisable to use the HbA1c criteria for diagnosis of DM [2,10], being considered as generally equally appropriate as the Fasting Plasma Glucose (FPG) value and Oral Glucose Tolerance Test (OGTT). In contrast to the FPG and OGTT diagnostics, the HbA1c test does not require patient preparation because it provides a long-term indicator of mean glycemia, which is robust against glucose peaks in blood associated to many physiological processes [11]. On the other hand, although short-term glycemic variability or hypoglycemic events are increasingly addressed using Continuous Glucose Monitoring (CGM) metrics [12], which also provides the Glucose Management Indicator (GMI), the hemoglobin A1c is still considered as the gold standard in medium/long-term glycemic control due to the widely demonstrated relationship between hemoglobin A1c and diabetes-related microvascular complications.

Unfortunately, conventional methods for instantaneous BGL measurement and HbA1c test use invasive technologies (blood samples required) based on consumables, resulting in painful experiences, time-consuming procedures and heavy financial burdens. Besides this, currently available point-of-care (POC) instruments for rapid and outpatient HbA1c diagnostic generally offer poor performance compared with laboratory-based methods [13], resulting in a limited global availability and accessibility of adequate assays, especially in low- and middle-income countries. All these drawbacks explain the high rate of undiagnosed cases and the poor patients' compliance to follow completely medical recommendations in DM care [14].

The clear need to improve current available technologies for DM management has attracted tremendous attention in the last decades and the use of non-invasive technologies promises to be a life-changing factor with huge benefits in DM care. This technology evolution resulted in several trends aiming at the non-invasive BGL monitoring [15–18], but, as far as the authors are aware, just few proposals, based on Raman spectroscopy, piezoelectric detection, and optical methods, have been tested in-vitro [19–21] and in-vivo [22,23] to address the non-invasive assessment of medium/long-term sustained glycemia, compatible with the HbA1c test. In any case, no single technology has been successfully implemented in medical proceedings because the standards for an ideal sensor have not yet been achieved, leaving this issue still open.

In particular, the millimeter-wave (mm-wave) spectroscopy (30 GHz - 300 GHz) has shown great potential for the non-invasive sensing of BGL, with in-vivo blood glucose monitoring in animal models reported using Ka-band frequencies (27 GHz - 40 GHz) [24,25], and the detection of glucose spikes in humans during an intravenous glucose tolerance test using transmission measurements at 60 GHz [26].

In this article, a novel non-invasive spectroscopic approach, based on W-band spectroscopy (75 GHz – 110 GHz), is proposed and tested within a first clinical study for non-invasive detection and monitoring of sustained hyperglycemia, typically associated with DM. This non-invasive approach has been already successfully tested on animal models [27–29], showing great capability for in-vivo detection and monitoring of sustained glycemia. A pilot diagnostic validation study was conducted on a group of thirty volunteers, involving type 1 diabetes mellitus (T1DM) patients and healthy subjects, to assess the relationship between their HbA1c concentrations and the non-invasive measure provided by the proposed spectroscopic technique.

2. Experimental setup: millimeter-wave spectroscopy system

The spectroscopic approach proposed for the non-invasive assessment of sustained hyperglycemia is based on a mm-wave spectrometer spanning the whole W-band from 75 GHz to 111 GHz. The W-band offers a good balance between penetration depth in tissue, interaction volume, and signal-to-noise ratio [15,26,30]; and the non-invasive instrument can be made very compact and mass-produced using semiconductor processes, which is very promising for commercial development. The spectrometer uses a frequency multiplication chain to reach the W-band frequencies from a Ku-band generator (12 GHz - 18 GHz). The mm-wave spectroscopy system

used for the non-invasive assessment is thoroughly described in reference [23] and it is shortly described below.

A synthesized signal generator APSIN20G (AnaPico, Zurich, Switzerland) is used to generate a frequency sweep from 12.5 GHz to 18.5 in steps of 0.125 GHz (49 frequencies equally distributed). The input frequency from the signal generator is multiplied by a factor of $\times 6$ by an AFM6 Active Frequency Multiplier (Radiometer Physics GmbH, Meckenheim, Germany), resulting in a frequency sweep from 75 GHz to 111 GHz in steps of 0.75 GHz. The output of the AFM6 frequency multiplier is fed into a dual directional coupler with three output ports corresponding to the reference channel, the reflection channel, and a thru branch. The reference and reflection ports are directly connected to two HMR6 subharmonic mixer receivers (Radiometer Physics GmbH, Meckenheim, Germany), and the signal of the thru branch is directed towards a skin fold of the study subjects by using a waveguide probe. Then, the signal that traveled through the skin fold (transmission) is amplified by a W-band low noise amplifier (Radiometer Physics GmbH, Meckenheim, Germany) and detected by a third HMR6 receiver. The local oscillator signal for the HMR units is also generated by a slightly detuned output of the APSIN20G synthesizer. The probes are previously aligned straight cuts of a WR10 waveguide tapered on the end directly in contact with the skin fold of the study subjects. The outputs of the subharmonic mixer receivers deliver the detected signals (reference, reflection and transmission) at an intermediate frequency of $f_{IF} = 1.8$ MHz that are sampled using a data acquisition unit (Handyscope HS4-10, TiePie engineering, Sneek, Netherlands) at a sampling rate of 10 MHz. Finally, amplitude and phase parameters of the digitized signals are acquired in real time through a multi-channel lock-in amplifier implemented in LabVIEW software [31].

Let us note that, as differently from the system described in [23], the amplification stage (low noise amplifier) used for the transmission signal (after the skin fold) was added for human measurements (not included in setup used on animal models [27–29]) to compensate for human tissue absorption. Also in line with our previous work using animal models [27–29], the pilot diagnostic validation study used the amplitude of the detected signal after the skin fold of the study subjects (transmission mode), since this parameter showed great potential for detection of sustained hyperglycemic states in our previous experimental tests on mice.

3. Pilot validation diagnostic test: sample and measurement protocol

The pilot diagnostic assessment was carried out in collaboration with the Instituto de Investigación Sanitaria of the Hospital Fundación Jiménez Díaz (Servicio de Endocrinología y Nutrición) under a protocol study (Protocol code: FJD-ESPEC-DM-17-01) approved by the hospital's Ethics Committee (CEI/CEIm-FJD). The clinical procedures were conducted in accordance with the ethical principles regarding human experimentation established in the Declaration of Helsinki developed by the World Medical Association, the guidelines for Good Clinical Practice of the International Committee of Harmonization, and the Spanish Law 14/2017, of 3 July, on Biomedical Research. All the information related to the participants was processed according to current personal data protection regulations.

Thirty subjects were recruited under informed consent after having understood about the objectives and protocol of the pilot diagnostic validation study. Of the thirty subjects, twenty subjects had been previously diagnosed of T1DM, while the remaining ten had never been diagnosed with DM (healthy subjects). Two different types of metabolic control were equivalently included within the group with T1DM: patients with good metabolic control ($HbA1c < 7\%$) and patients with poor metabolic control ($HbA1c > 8\%$), according to their clinical records at the time of the recruitment. The sample population related to the pilot diagnostic validation study is summarized in Table 1.

Table 1. Sample population evaluated.

Condition	HbA1c	Quantity
Control subjects without DM	-	10
T1DM in good metabolic control	< 7%	10
T1DM in poor metabolic control	> 8%	10

The mean age of the study subjects was 47 years with a standard deviation of 14.8 years (ranging from 25 to 79), the gender distribution was 60% women and 40% men, and the mean body mass index (BMI) was 23.89 kg/m² with a standard deviation of 3.44 kg/m².

Only T1DM was considered in the pilot diagnostic validation test since it affects homogeneously across the population, in contrast to type 2 DM which is more frequent in older adults. The diagnostic criteria considered for inclusion of T1DM patients in the study were the following [2]:

- a. The presence of cardinal signs of diabetes such as polyuria, polydipsia, and unexplained weight loss, with plasma glucose ≥ 200 mg/dl or diagnosis of diabetic ketoacidosis.
- b. Fasting plasma glucose (≥ 8 hrs) ≥ 126 mg/dl.
- c. Plasma glucose ≥ 200 mg/dl after two hours in the oral glucose tolerance test (75g of glucose).
- d. HbA1c $\geq 6.5\%$ (according to the National Glycohemoglobin Standardized Program and standardized by the Diabetes Control and Complications Trial).

The exclusion criteria were diagnosis of type 2 DM or monogenic diabetes, pregnancy, severe renal or hepatic insufficiency, poor short-term prognosis (< 6 months), treatment with glucocorticoids or immunosuppressive medications, or those who had been hospitalized in the past three months.

The pilot diagnostic validation test was designed as a prospective study consisting of 3 visits recorded within a period of 7 months, with 3 months between visits 1 and 2, and 4 months between visits 2 and 3. The general outline of the pilot diagnostic validation test is summarized

Table 2. Summary of the pilot diagnostic validation test.

Visit 0	Recruitment (compliance assessment of selection criteria). Supplementing informed consent forms.
Visit 1	Non-invasive assessment using the mm-wave spectrometer. HbA1c test. Blood analysis (clinical biochemistry). Collection of physiological and anthropometric variables.
3 months	
Visit 2	Non-invasive assessment using the mm-wave spectrometer. HbA1c test. Blood analysis (clinical biochemistry). Collection of physiological and anthropometric variables.
4 months	
Visit 3	Non-invasive assessment using the mm-wave spectrometer. HbA1c test. Blood analysis (clinical biochemistry). Collection of physiological and anthropometric variables.

in Table 2 indicating the tests carried out at each visit. At each visit, the study subjects were non-invasively assessed to obtain their transmission W-band spectral profiles, and a HbA1c test was done to measure their glycated hemoglobin values. HbA1c was measured by a D-100 hemoglobin testing system that uses smart High-Performance Liquid Chromatography (HPLC) technology [32], and results were standardized to those of the diabetes control and complications trial [33]. In addition, a blood analysis (clinical biochemistry), including clinical biomarkers such as triglycerides, cholesterol, creatinine, liver transaminases, and hemoglobin content, was collected at each visit, together with other physiological and anthropometric variables, such as age, weight, height, body mass index, skin-fold thickness and blood pressure.

The visits were scheduled at regular medical appointments, and the study subjects were non-invasively assessed using the mm-wave spectrometer under medical supervision at the Hospital Universitario Fundación Jiménez Díaz.

As shown in Fig. 1, the non-invasive assessment was performed on a skin fold located in the first interdigital space (between the thumb and index finger) of the right hand. The first interdigital space was chosen because it is highly vascularized [34,35], the skin is relatively thin, and it is a comfortable location to perform the measurement. In order to reduce arm movement of the study subjects and to ensure their comfort during the non-invasive assessment, an elbow support was incorporated to the structure of the W-band spectroscopic instrument (Fig. 1). The complete measurement process, covering the 49 frequencies equally distributed over the W-band takes around one minute. The measurement time is mainly limited by the control electronics (signal generator and LabVIEW program for signal acquisition and processing) rather than the W-band spectroscopic instrument.

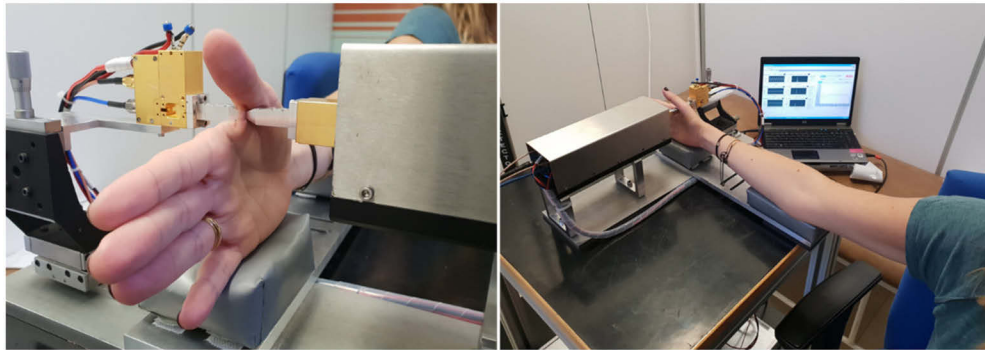


Fig. 1. Photographs taken during the non-invasive assessment of sustained hyperglycemia using the mm-wave spectrometer installed at the Hospital Fundación Jiménez Díaz.

Let us note that no special indications concerning the non-invasive assessment were given to the participants, and all the patients always continued their medical treatments.

The output power of the signal generator was calibrated in the laboratory prior to its installation in the hospital to obtain a flat frequency response at the transmission port (around $250 \text{ mV} \pm 5 \text{ mV}$). Such calibration was carried out using a RF attenuator (17 dB) at the input of the AFM6 frequency multiplier to compensate for human tissue absorption and ensure a good dynamic range. A separation of 1.8 mm between the probes of the W-band spectroscopic instrument was fixed for calibration tests and to hold the skin fold of the study subjects during the non-invasive assessments at the visits. It should be emphasized that in contrast to other non-invasive techniques no pre-treatment of the hand or skin (washing, treatment with alcohol, etc.) has been undertaken and the skin-waveguide contact pressure had insignificant impact on the results, as long as contact was established.

4. Data processing and statistical methods

The non-invasive assessment of study subjects consists of an analysis of the acquired transmission spectral profiles at 49 frequency points equally distributed over the W-band, as described above, using Functional Data Analysis (FDA) [36,37] for data post-processing. The FDA is a branch of Statistics which is increasingly promoted in the field of applied spectroscopy since it allows to analyze spectral data by considering the continuous intrinsic nature that generally characterizes spectroscopic measurements, among other advantages [27,38,39]. In this sense, and prior to the functional regression analysis, the collected spectra (raw data measured by the spectroscopy instrument) given by $\{x_{ijk} : i = 1, \dots, N; j = 1, \dots, J_i; k = 1, \dots, m\}$, where N is the total number of subjects, J_i the number of visits recorded for the i -th subject, and m the number of measured frequency points, were used to approximate the continuous trajectories of the transmission W-band spectral profiles (sample curves) $\{x_{ij}(f) : i = 1, \dots, N; j = 1, \dots, J_i; f \in F\}$, where F is the frequency interval (75 GHz - 111 GHz). In other words, we want to approximate a set of sample curves (as observations of a frequency-dependent function) that best describe the measured spectra represented by vectors, in which the entries correspond to the frequencies measured during the spectral interrogation. To this end, it is assumed that the sample curves belong to a finite-dimensional space spanned by a basis $\{\phi_1(f), \phi_2(f), \dots, \phi_p(f)\}$, so that they can be expressed as follows:

$$x_{ij}(f) = \sum_{l=1}^p a_{il} \phi_l(f). \quad (1)$$

Then, considering that the collected spectra were measured with error (noise), and the underlying frequency-dependent function is smooth, we use Penalized splines (P-splines) [40] for the estimation of the sample curves. The P-splines were estimated using cubic B-splines as the basis functions $\phi_l(f)$ and with the basis coefficients a_{il} computed by penalized least squares [41]. In the estimation of the P-splines a penalty term is added in the least square's equation with a smoothing parameter λ that defines the degree of smoothness in the approach. The λ value was chosen using the Generalized Cross Validation (GCV) method [41,42] such that it minimizes the averaged mean square error between the approximated sample curves and the collected spectra.

In general, the estimated sample curves are analyzed under two different approaches: non-longitudinal, i.e., three functional regression models are estimated (one for each visit), and longitudinal, i.e., all the sample curves are modelled simultaneously considering that the subjects were assessed repeatedly over a given period (in our case 3 visits for 7 months).

For the non-longitudinal analysis of the sample curves (taking each visit separately), the relationship between the transmission W-band spectral profile of the study subjects and their HbA1c concentrations was modelled by scalar-on-function linear regression [36,43], with the sample curves $x_i(f)$ as the functional predictor and the HbA1c value as the scalar response variable y_i . The scalar-on-function regression model is then formulated as follows:

$$y_i = \alpha + \int_F x_i(f) \beta(f) df + \epsilon_i, \quad (2)$$

where α is a scalar intercept, $\beta(f)$ is the functional parameter of the regression model, and ϵ_i are independent and identically distributed (iid) errors with zero mean. More complex methods, involving supervised and unsupervised approaches, considered for longitudinal analysis are briefly introduced below.

A Longitudinal Functional Principal Component Analysis (LFPCA) [44] was performed to explore variability within the transmission W-band spectral profiles. The LFPCA is an unsupervised statistical approach, based on the principle of the classical PCA, providing an easy way to evaluate the Principal Components (PCs) of the variance contained in the sample curves. The LFPCA takes into account the variability between different subjects (inter-subject variation),

which is the standard variance considered for classical regression analysis, and the variability within the sample curves coming from a same subject (intra-subject variation), i.e., how the transmission W-band spectral profile of a specific subject changes over visits. Both sources of variation (inter- and intra-subject) are decomposed in a set of PCs (as in classical PCA), and the scores of the PCs can be used to look for natural clusters and to study longitudinal relations to clinical variables of interest. In the LFPCA formulation, the longitudinal sample curves are modelled by a functional random intercept and random slope model which is estimated using different approaches, such as bivariate smoothing with P-splines, multiple linear regression with interactions terms, spectral decomposition, Karhunen-Loève expansion, and best linear unbiased prediction. For more details about the LFPCA theory and estimation, please refer to [44].

Finally, the relationship between the transmission W-band spectral profile of study subjects and their HbA1c concentrations was also modelled by performing a longitudinal functional regression analysis (supervised approach). To this end, two different longitudinal scalar-on-function regression models were considered: Longitudinal Penalized Functional Regression (LPFR) [45] and Longitudinal Functional Principal Component Regression (LFPCR) [46]. In the LPFR, the longitudinal sample curves are modelled similarly to Eq. (2), but an additional term is added to incorporate the intra-subject variability by scalar random-effects intercepts. In contrast, the LFPCR model is based on LFPCA for variance decomposition of the longitudinal sample curves, using the estimated scores from inter- and intra-subject variation as predictors.

All processing and analysis of the collected spectral data was made through the statistical free software R project [47,48]. The processing of collected spectral data and the approximation of the sample curves were performed using the “*fda*” package [49] and other codes developed by the authors. The scalar-on-function linear regression was implemented using the “*fRegress*” function also available in the “*fda*” package. The LFPCA [44], LPFR [45] and LFPCR [46] were performed using the R codes provided by the authors in their corresponding publications.

5. Results and discussion

As described above, a sample population of 30 subjects was recruited and non-invasively assessed using a mm-wave spectrometer at three different visits. The collected spectral data included the W-band spectral profiles (transmission amplitude measurement) from 10 T1DM with poor metabolic control (HbA1c > 8%), 10 T1DM with good metabolic control (HbA1c < 7%), and 10 healthy subjects without DM. After the 3 visits, a total of 83 spectral profiles were collected from the study subjects (30 for visit 1, 27 for visit 2 and 26 for visit 3) along with the corresponding HbA1c values. The collected spectral data were used to approximate the sample curves shown in Fig. 2 by P-splines defined on 35 equally spaced knots over the W-band, with the smoothing parameter $\lambda = 0.3$ chosen by the GCV method. Let us observe here that working with cubic B-spline basis, the basis size (p value introduced in Eq. (1)) equals the number of knots plus 2.

The analysis of the sample curves shown in Fig. 2 is addressed using both unsupervised and supervised approaches: first, the variability within the longitudinal sample curves is explored considering inter- and intra-subject variations, then, a regression analysis is performed to study the relationship between the transmission W-band spectral profiles and the HbA1c concentrations using non-longitudinal (scalar-on-function linear regression) and longitudinal (LPFR and LFPCR) methods. The main results obtained from both non-longitudinal and longitudinal analysis are shown below.

5.1. Blind detection of hyperglycemia metabolism

The LFPCA was performed on the sample curves estimating the first four Principal Components (PCs). A total of 95.8% of the variability in the W-band spectral profiles was captured, with a 54.8% related to the inter-subject variation and the remaining 41% related to the intra-subject variation. Then, the estimated scores from inter-subject variation were used to look for possible

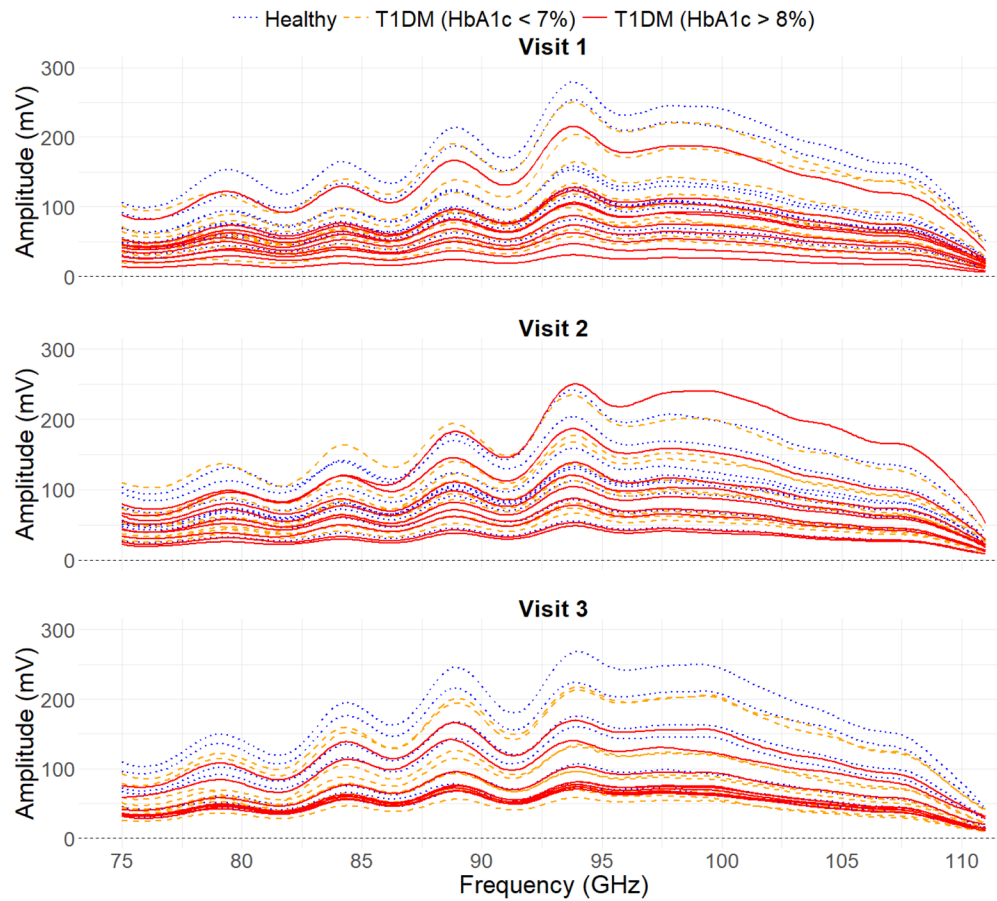


Fig. 2. Estimated sample curves from amplitude spectra of the transmitted wave through the skin fold of study subjects: healthy subjects that had never been diagnosed with DM (dotted blue lines), T1DM patients in good metabolic control (dashed orange lines), and T1DM in poor metabolic control (solid red lines).

natural clusters within the study subjects that might be related to sustained hyperglycemia, typically associated with DM. The mean HbA1c values of study subjects from their recorded visits are shown at the top of Fig. 3, with a boxplot displayed for each group: control subjects without DM (blue), T1DM patients with poor metabolic control (red), and T1DM patients with good metabolic control (orange). Similarly, at the bottom of this figure the scores corresponding to the first PC corresponding to the inter-subject variation (44.7% of explained variance) from the spectroscopic data are shown.

As depicted in Fig. 3, the scores estimated from the study subjects using the mm-wave spectrometer show similar dispersion as those observed for their averaged HbA1c values, allowing potential discrimination between healthy subjects and T1DM patients using the proposed diagnostic. This blind discrimination achieved from an unsupervised experiment suggests that our non-invasive instrument is able to detect changes in the metabolism of the study subjects associated to T1DM.

On the other hand, no significant differences were observed between T1DM patients with poor metabolic control ($\text{HbA1c} > 8\%$) and those with good metabolic control ($\text{HbA1c} < 7\%$). This could be expected because the patients were not under special treatments or diet regimens; instead,

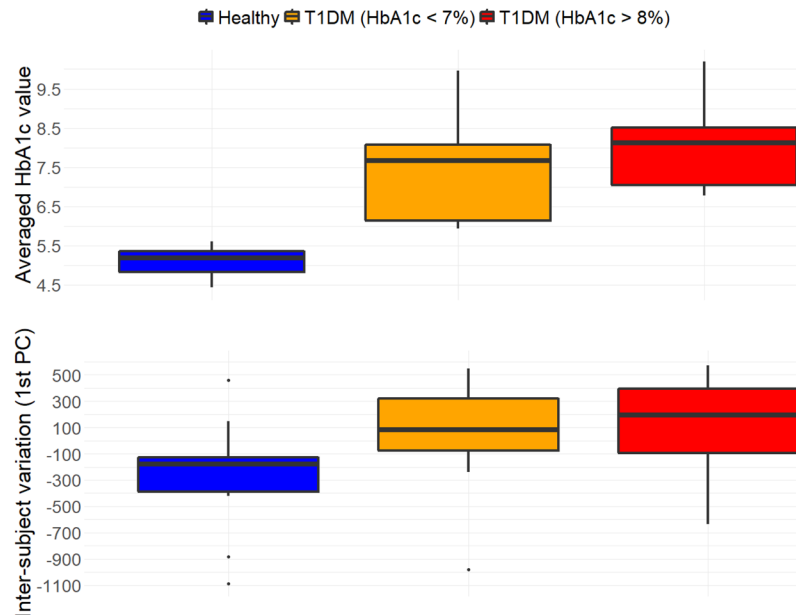


Fig. 3. Boxplots by groups: glycemic control of healthy subjects (blue), T1DM patients with poor metabolic control - HbA1c < 7% (red), and T1DM with good metabolic control - HbA1c > 7% (orange), showing the mean HbA1c values of study subjects from their recorded visits (at the top), and their estimated scores corresponding to the 1st PC from inter-subject variation (at the bottom) from the spectroscopic data.

they were told to stick to their regular medical treatments. In this sense it is important to note that such distinction (poor or good metabolic control) was assigned according to their clinical records at the recruitment stage (and thus, before the measurements), so the actual metabolic control of the patients may vary during the duration of the clinical trial.

Similarly, the estimated scores from intra-subject variation, capturing a total of 41% of the variability within longitudinal sample curves, were used to study possible relations between the evolution of the transmission W-band spectral profiles among visits and other collected clinical variables, such as hemoglobin content, triglycerides, cholesterol, creatinine, liver transaminases, age, weight, height, body mass index, skin-fold thickness and blood pressure. The variability of the instrument itself has been inherently considered, but experience showed that the instrument's variability is very small as compared to other biological factors. Then, the Pearson's correlation coefficients of the clinical variables against the first four PCs were estimated by pairs, but no significant correlations were found.

The results shown in Fig. 3 support that, similarly to the diagnosis of DM provided by the HbA1c test ($\text{HbA1c} \geq 6.5\%$), the W-band spectroscopy has potential for non-invasive detection of T1DM. However, the major aim of the pilot diagnostic validation study was to test the applicability of W-band spectroscopy for monitoring sustained hyperglycemia. Such a task is addressed below using both non-longitudinal and longitudinal scalar-on-function linear regression on HbA1c values.

5.2. Linear relationship with glycated hemoglobin

The applicability of W-band spectroscopy for monitoring sustained hyperglycemia is tested by performing scalar-on-function linear regression on the HbA1c concentrations from each visit with the corresponding estimated sample curves as functional predictor. In Fig. 4 is shown the

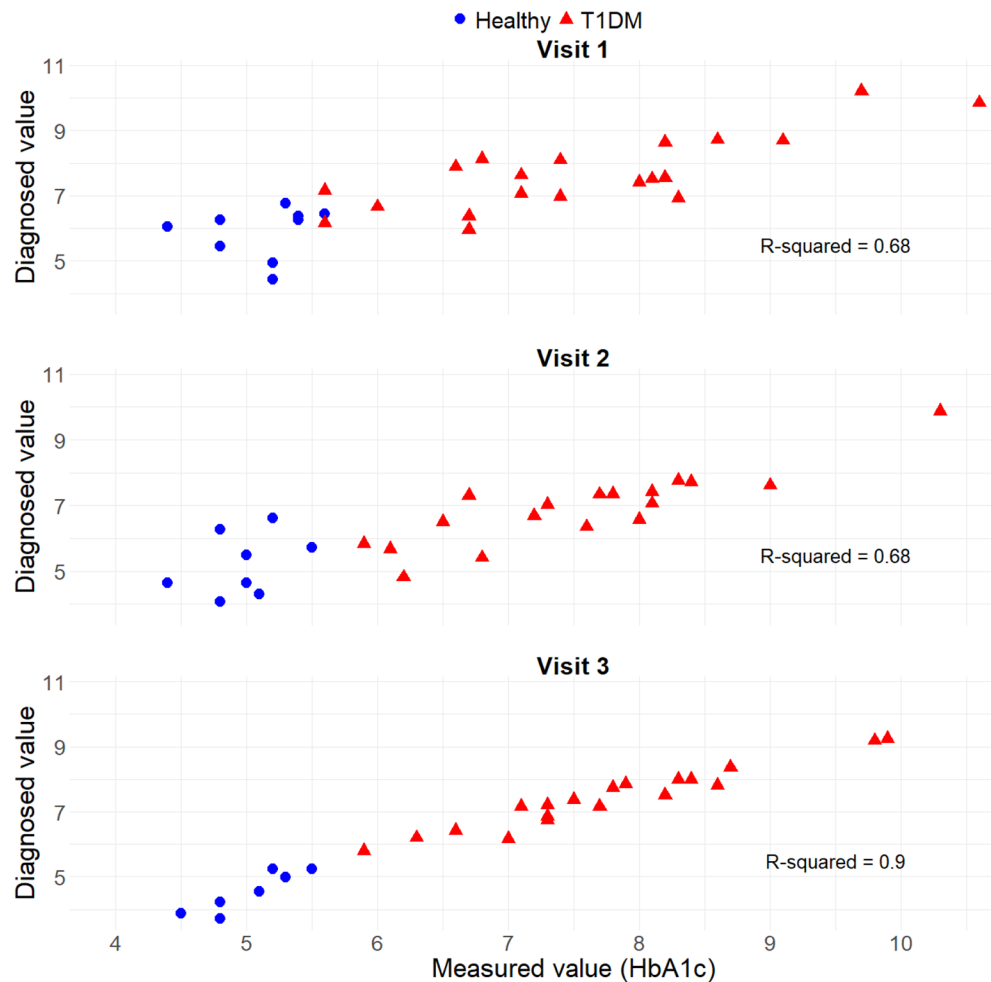


Fig. 4. Scatterplots showing the HbA1c concentrations versus predicted values by the spectrometer diagnostic at each visit. Healthy subjects without DM and T1DM patients are identified by blue circles and red triangles, respectively.

best fit achieved for HbA1c concentrations at each visit. In order to get the best fit, the functional parameter of the regression models was estimated considering different number of basis functions (cubic B-spline basis functions) for each visit (visit 1 = 19, visit 2 = 21, and visit 3 = 18). It should be mentioned that two sample curves had to be excluded from the regression analysis because their HbA1c values were missing, leaving a total of 81 observations in the regression analysis: 29 subjects for visit 1 and 26 subjects for visits 2 and 3.

As displayed in Fig. 4, the achieved proportion of variance explained of the HbA1c concentrations by the functional regression models (R-squared value) varies among visits, with the best fit (R-squared = 0.9) obtained on the last visit. The significant difference of the R-squared value observed between the last visit and the first two visits implies that there are uncontrolled factors involved in the non-invasive assessment that strongly affects the HbA1c value prediction. Such variation in the obtained results might be caused by changes in the calibration of the spectroscopic instrument (recalibration between visits with 3 months in between), or changes in metabolism of the study subjects between visits (within-subject variation among visits). Besides this, other

interferences in the non-invasive assessment such as physiological, environmental or instrumental noise cannot be dismissed. Nevertheless, the results of the last visit clearly indicate that a good linear relation can be achieved between the transmission W-band spectral profile and the HbA1c concentration.

Finally, longitudinal functional regression was considered to assess whether the within-subject variation is related to the sustained hyperglycemia condition, i.e., all the 81 observations were modelled simultaneously by the regression model. Two different longitudinal scalar-on-function regression models (LPFR and LFPCR) were estimated to measure the relation between the transmission W-band spectral profiles and the HbA1c value. The main difference between the LPFR and the LFPCR approaches is that the first one directly models the sample curves by estimating a functional parameter and the second one (score-based approach) is based on the variance decomposition estimated by the LFPCA (inter- and intra-subject variation). The LFPCR model was estimated by considering the first three PCs with a cumulative explained variance of 95.4% (53.4% of inter-subject variation and 42% of intra-subject variation).

In Fig. 5 we show a scatter plot of the measured HbA1c values against the predicted values by the fitted LPFR model. As it can be seen, the LPFR model provides an excellent fit of the HbA1c values measured from all visits, with a R-squared value of 0.98 and a root mean square error (RMSE) equal to 0.24. We would like to stress that this result has been achieved by considering only the W-band spectral profile as a functional predictor variable, i.e., no other clinical variables were considered or used in the model estimation.

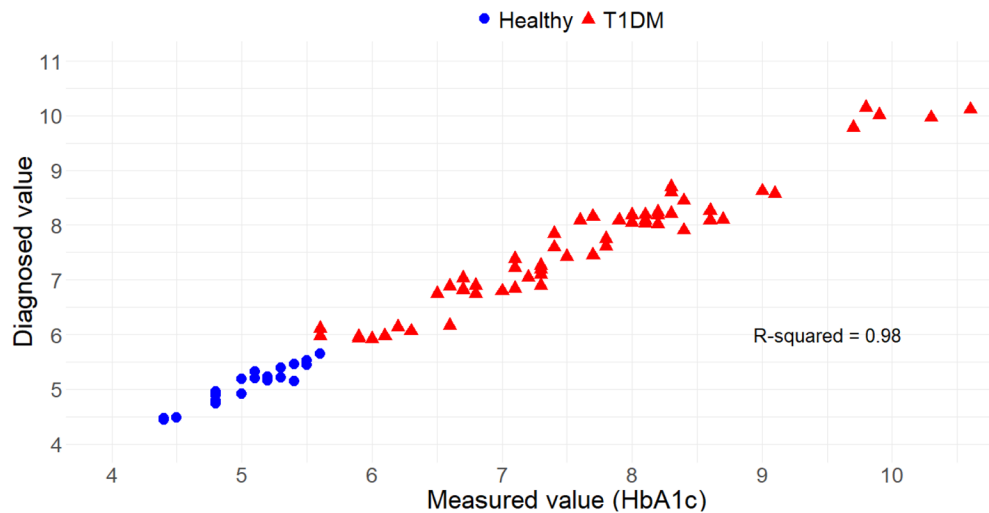


Fig. 5. Scatterplot showing the HbA1c values versus provided values by the spectrometer diagnostic. Healthy subjects without DM and T1DM patients are identified by blue circles and red triangles, respectively.

The improved R-squared value indicate that intra-subject variation is also partly related to the hyperglycemia condition, significantly improving the regression analysis for all the study subjects. These are very promising results and indicate that the transmission W-band spectral profile acquired by the mm-wave spectrometer has great potential for non-invasive T1DM monitoring, being strongly related to the biomarker of HbA1c test. Such relation was validated by the LFPCR model estimated from the same sample curves. The comparison of the performance of both models can be seen in Table 3. An excellent fit is also achieved by the LFPCR model (scores-based approach) with very similar results, achieving a R-squared equal to 0.98 and a RMSE equal to 0.23. The consistent results obtained from the two different estimation methods

clearly demonstrate the potential of the proposed approach. The excellent linear relationship between the transmission W-band spectral profile and the HbA1c concentration supports the usefulness of W-band frequencies for diagnosis and monitoring of sustained hyperglycemia typically associated with DM. It should also be emphasized here that the healthy subjects can be fully distinguished from the T1DM patients. The only slight overlap can be observed at a HbA1c value of 5.5 where the predicted values are around 6 as compared to the measured value and the results for the healthy subject and T1DM patients overlap.

Table 3. Results obtained from the fitted longitudinal models.

Model	R-squared	RMSE
LPFR	0.98	0.24
LFPCR (scores-based)	0.98	0.23

6. Conclusions

In this study, a clinical diagnostic validation was carried out testing the applicability of W-band spectroscopy for a novel non-invasive diagnostic tool monitoring sustained hyperglycemia, typically associated with DM. A group of 30 subjects (20 T1DM and 10 control subjects) was non-invasively assessed using a mm-wave spectrometer at three different visits during a period of 7 months. A non-invasive assessment was performed directly on a skin fold located in the first interdigital space of the right hand. Besides the non-invasive assessment, subjects underwent a HbA1c test (glycated hemoglobin) at each visit to monitor their levels of sustained hyperglycemia. A set of 83 sample curves (P-splines), approximated from raw data, was used to analyze the W-band spectral profiles of study subjects within the FDA framework.

First, the variability in the sample curves was explored by performing a LFPCA (unsupervised analysis) to look for natural clusters within the study subjects according to their condition. A key research finding was that the first component capturing most of the inter-subject variation (with a 44.7% of explained variance) provided potential discrimination (blindly) between diabetic and non-diabetic patients. Such discrimination shows that the inter-subject variation within the W-band spectral profiles is mostly related to the presence of T1DM in subjects. This suggest that the non-invasive approach can be used for instantaneous diagnosis of DM.

A second analysis was conducted on sample curves to test the relationship between the transmission W-band spectral profiles and the HbA1c values by scalar-on-function linear regression models, considering non-longitudinal and longitudinal methods. Firstly, the sample curves approximated from each visit were analyzed separately (non-longitudinally) performing scalar-on-function linear regression. The obtained results provided evidence that a good linear relation can be achieved between the transmission W-band spectral profile and the HbA1c concentrations, but the model fit (R-squared) is considerably affected among visits decreasing from 0.9 to 0.68. This suggest that there are uncontrolled factors introducing changes in the W-band spectral profile of subjects among visits (intra-subject variation) that need to be considered in the regression analysis.

Finally, in order to study whether the intra-subject variation is partly related to the sustained hyperglycemia condition, such intra-subject variation was included in the regression analysis by estimating two longitudinal regression models, LFPCR and LPFR, which are based on two different approaches. The fitted models provided excellent results achieving a R-squared equal to 0.98 with a RMSE equal to 0.23.

All the obtained results are very promising since they demonstrate that the measured W-band spectral profiles (transmission amplitude measurement) can be employed as an excellent indicator for sustained glycemia in humans, closely related to the HbA1c value. The blind identification of

subjects exhibiting diabetic and non-diabetic states, together with the strong relationship between the HbA1c values and the measured spectra, are clear experimental evidence supporting that W-band transmission spectroscopy can potentially be used for in-vivo and non-invasive detection and monitoring of sustained hyperglycemia in humans.

The results presented in this paper might pave the way for the development of a novel generation of instruments for non-invasive diagnosis and monitoring of sustained hyperglycemia. Then, the next step in this research line is to validate the achieved relation between the W-band spectral transmission assessment and the HbA1c concentration through the realization of more clinical studies involving larger number of subjects. A more in-depth study will also be carried out to specifically identify factors influencing the intra-subject variability to improve the consistency of the non-longitudinal linear regression analysis. Finally, it is also intended those measurements in the following clinical study will be performed fusing a preliminary Point-of-Care (POC) version of the W-band spectroscopic instrument.

Funding. Instituto de Salud Carlos III (DTS17/00135); Comunidad de Madrid (EPUC3M26).

Acknowledgements. Aldo Moreno-Oyervides thanks the Consejo Nacional de Ciencia y Tecnología de México (CONACYT) for financially supporting his doctoral education (PhD. Grant) associated to this work.

This work has been supported by the Madrid Government (Comunidad de Madrid-Spain) under the Multiannual Agreement with UC3M in the line of Excellence of University Professors (EPUC3M26), and in the context of the V PRICIT (Regional Programme of Research and Technological Innovation).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the corresponding author upon reasonable request.

References

1. P. Saeedi, I. Petersohn, P. Salpea, B. Malanda, S. Karuranga, N. Unwin, S. Colagiuri, L. Guariguata, A. A. Motala, K. Ogurtsova, J. E. Shaw, D. Bright, and R. Williams, "Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the International Diabetes Federation Diabetes Atlas, 9th edition," *Diabetes Res. Clin. Pract.* **157**, 107843 (2019).
2. American Diabetes Association, "Standards of Medical Care in Diabetes 2021," *Diabetes Care* **44**(Supplement 1), (2021).
3. R. Singh, A. Barden, T. Mori, and L. Beilin, "Advanced glycation end-products: a review," *Diabetologia* **44**(2), 129–146 (2001).
4. V. P. Singh, A. Bali, N. Singh, and A. S. Jaggi, "Advanced glycation end products and diabetic complications," *Korean J. Physiol. Pharmacol.* **18**(1), 1–14 (2014).
5. A. W. Stitt, A. J. Jenkins, and M. E. Cooper, "Advanced glycation end products and diabetic complications," *Expert Opin. Investig. Drugs* **11**(9), 1205–1223 (2002).
6. K. Papatheodorou, N. Papanas, M. Banach, D. Papazoglou, and M. Edmonds, "Complications of diabetes 2016," *J. Diabetes Res.* **2016**, 6989453 (2016).
7. M. Brownlee, "The pathobiology of diabetic complications: a unifying mechanism," *Diabetes* **54**(6), 1615–1625 (2005).
8. C. Weykamp, "HbA1c: a review of analytical and clinical aspects," *Ann. Lab. Med.* **33**(6), 393–400 (2013).
9. C. D. Saudek and J. C. Brick, "The clinical use of hemoglobin A1c," *J. Diabetes Sci. Technol.* **3**(4), 629–634 (2009).
10. N. Alqahtani, W. A. G. Khan, M. H. Alhumaidi, and Y. A. A. R. Ahmed, "Use of glycated hemoglobin in the diagnosis of diabetes mellitus and pre-diabetes and role of fasting plasma glucose, oral glucose tolerance test," *Int. J. Prev. Med.* **4**(9), 1025 (2013).
11. R. Derr, E. Garrett, G. A. Stacy, and C. D. Saudek, "Is HbA1c affected by glycemic instability," *Diabetes Care* **26**(10), 2728–2733 (2003).
12. H. Chehregosha, M. E. Khamseh, M. Malek, F. Hosseinpahan, and F. Ismail-Beigi, "A view beyond HbA1c: role of continuous glucose monitoring," *Diabetes Ther.* **10**(3), 853–863 (2019).
13. E. Lenters-Westra and R. J. Slingerland, "Six of eight hemoglobin A1c point-of-care instruments do not meet the general accepted analytical performance criteria," *Clin. Chem. (Washington, DC, U. S.)* **56**(1), 44–52 (2010).
14. J. E. F. Ward, B. A. Stetson, and S. P. L. Mokshagundam, "Patient perspectives on self-monitoring of blood glucose: perceived recommendations, behaviors and barriers in a clinic sample of adults with type 2 diabetes," *J. Diabetes Metab. Disord.* **14**(1), 43 (2015).
15. W. V. Gonzales, A. T. Mobashsher, and A. Abbosh, "The progress of glucose monitoring—A review of invasive to minimally and non-invasive techniques, devices and sensors," *Sensors* **19**(4), 800 (2019).
16. A. Nawaz, P. Øhlckers, S. Sælid, M. Jacobsen, and M. Nadeem Akram, "Review: non-invasive continuous blood glucose measurement techniques," *J. Bioinform. Diabetes* **1**(3), 1–27 (2016).

17. M. Shokrehodaie and S. Quinones, "Review of non-invasive glucose sensing techniques: optical, electrical and breath acetone," *Sensors* **20**(5), 1251 (2020).
18. M. Shokrehodaie, D. P. Cistola, R. C. Roberts, and S. Quinones, "Non-invasive glucose monitoring using optical sensor and machine learning techniques for diabetes applications," *IEEE Access* **9**, 73029–73045 (2021).
19. J. Sundararaj, V. Palanisamy, and M. Sandeep, "A novel and proven system for non-invasive blood glucose monitoring using HbA1C," *Asian J. Appl. Sci.* **2**(3), 253–274 (2009).
20. S. Mandal and M. O. Manasreh, "An in-vitro optical sensor designed to estimate glycated hemoglobin levels," *Sensors* **18**(4), 1084 (2018).
21. I. Barman, N. C. Dingari, J. W. Kang, G. L. Horowitz, R. R. Dasari, and M. S. Feld, "Raman spectroscopy-based sensitive and specific detection of glycated hemoglobin," *Anal. Chem.* **84**(5), 2474–2482 (2012).
22. E. Guevara, J. C. Torres-Galván, M. G. Ramírez-Elías, C. Luevano-Contreras, and F. J. González, "Use of Raman spectroscopy to screen diabetes mellitus with machine learning tools," *Biomed. Opt. Express* **9**(10), 4998–5010 (2018).
23. J. F. Villa-Manríquez, J. Castro-Ramos, F. Gutiérrez-Delgado, M. A. López-Pacheco, and A. E. Villanueva-Luna, "Raman spectroscopy and PCA-SVM as a non-invasive diagnostic tool to identify and classify qualitatively glycated hemoglobin levels in vivo," *J. Biophotonics* **10**(8), 1074–1079 (2017).
24. P. H. Siegel, Y. Lee, and V. Píkov, "Millimeter-wave non-invasive monitoring of glucose in anesthetized rats," in *International Conference on Infrared, Millimeter, and Terahertz Waves, IRMMW-THz* (IEEE Computer Society, 2014).
25. P. H. Siegel, W. Dai, R. A. Kloner, M. Csete, and V. Píkov, "First millimeter-wave animal in vivo measurements of L-Glucose and D-Glucose: Further steps towards a non-invasive glucometer," in *International Conference on Infrared, Millimeter, and Terahertz Waves, IRMMW-THz* (IEEE Computer Society, (2016), 2016-Novem.
26. S. Saha, H. Cano-Garcia, I. Sotiriou, O. Lipscombe, I. Gouzouasis, M. Koutsoupidou, G. Palikaras, R. Mackenzie, T. Reeve, P. Kosmas, and E. Kallos, "A glucose sensing system based on transmission measurements at millimetre waves using micro strip patch antennas," *Sci. Rep.* **7**(1), 1–11 (2017).
27. A. Moreno-Oyervides, M. C. Aguilera-Morillo, F. Larcher, V. Krozer, and P. Acedo, "Advanced statistical techniques for noninvasive hyperglycemic states detection in mice using millimeter-wave spectroscopy," *IEEE Trans. Terahertz Sci. Technol.* **10**(3), 237–245 (2020).
28. A. Moreno-Oyervides, P. Martín-Mateos, M. C. Aguilera-Morillo, G. Ulisse, M. C. Arriba, M. Durban, M. Del Rio, F. Larcher, V. Krozer, and P. Acedo, "Early, non-invasive sensing of sustained hyperglycemia in mice using millimeter-wave spectroscopy," *Sensors* **19**(15), 576 (2019).
29. P. Martín-Mateos, F. Dornuf, B. Duarte, B. Hils, A. Moreno-Oyervides, O. E. Bonilla-Manrique, F. Larcher, V. Krozer, and P. Acedo, "In-vivo, non-invasive detection of hyperglycemic states in animal models using mm-wave spectroscopy," *Sci. Rep.* **6**(1), 1–8 (2016).
30. N. K. Nikolova, "Microwave Biomedical Imaging," in *Wiley Encyclopedia of Electrical and Electronics Engineering* (John Wiley & Sons, Inc., (2014), pp. 1–22.
31. J. R. Lajara Vizcaíno and J. Sebastián Pelegrí, *LabVIEW: Entorno Gráfico de Programación*, 2nd ed. (MARCOMBO, S.A., 2011).
32. K. Lee, S. M. Kim, S. H. Jun, S. H. Song, K. U. Park, and J. Song, "Evaluation of Analytical Performance of the D-100 Hemoglobin Testing System for Hemoglobin A1c Assay," *J. Lab. Med. Qual. Assur.* **38**(2), 95–101 (2016).
33. D. M. Nathan, S. Genuth, L. Lachin, P. Cleary, O. Crofford, M. Davis, L. Rand, and C. Siebert, "The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus," *N. Engl. J. Med.* **329**(14), 977–986 (1993).
34. Y. Tsuruo, T. Ueyama, T. Ito, S. Nanjo, H. Gyoubu, K. Satoh, Y. Iida, and S. Nakai, "Persistent median artery in the hand: A report with a brief review of the literature," *Kaibogaku Zasshi* **81**(4), 242–252 (2006).
35. B. Strauch and W. de Moura, "Arterial system of the fingers," *J. Hand Surg. Am.* **15**(1), 148–154 (1990).
36. J. Ramsay and B. W. Silverman, *Functional Data Analysis*, 2nd ed. (Springer-Verlag, 2005).
37. J. O. Ramsay and B. W. Silverman, *Applied Functional Data Analysis*, 1st ed. (Springer-Verlag, 2002).
38. Z. Barati, I. Zakeri, and K. Pourrezaei, "Functional data analysis view of functional near infrared spectroscopy data," *J. Biomed. Opt.* **18**(11), 117007 (2013).
39. H. Sørensen, J. Goldsmith, and L. M. Sangalli, "An introduction with medical applications to functional data analysis," *Stat. Med.* **32**(30), 5222–5240 (2013).
40. P. H. C. Eilers and B. D. Marx, "Flexible smoothing with B-splines and penalties," *Stat. Sci.* **11**(2), 89–121 (1996).
41. A. M. Aguilera and M. C. Aguilera-Morillo, "Comparative study of different B-spline approaches for functional data," *Math. Comput. Model.* **58**(7–8), 1568–1579 (2013).
42. P. Craven and G. Wahba, "Smoothing noisy data with spline functions - Estimating the correct degree of smoothing by the method of generalized cross-validation," *Numer. Math.* **31**(4), 377–403 (1978).
43. P. T. Reiss, J. Goldsmith, H. L. Shang, and R. T. Ogden, "Methods for scalar-on-function regression," *Int. Stat. Rev.* **85**(2), 228–249 (2017).
44. S. Greven, C. Crainiceanu, B. Caffo, and D. Reich, "Longitudinal functional principal component analysis," *Electron. J. Statist.* **4**, 1022–1054 (2010).
45. J. Goldsmith, C. M. Crainiceanu, B. Caffo, and D. Reich, "Longitudinal penalized functional regression for cognitive outcomes on neuronal tract measurements," *J. R. Stat. Soc. Ser. C Appl. Stat.* **61**(3), 453–469 (2012).

46. J. Gertheiss, J. Goldsmith, C. Crainiceanu, and S. Greven, “Longitudinal scalar-on-functions regression with application to tractography data,” *Biostatistics* **14**(3), 447–461 (2013).
47. R Core Team, “R: a language and environment for statistical computing,” (2020).
48. J. O. Ramsay, G. Hooker, and S. Graves, *Functional Data Analysis with R and MATLAB* (Springer-Verlag, 2009).
49. J. O. Ramsay, S. Graves, and G. Hooker, “fda: functional data analysis,” (2020).